

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: August 22, 2003, 13:48:21; Search time 31.5 seconds

(without alignments)  
4636.982 Million cell updates/sec

Title: US-09-745-506-74

Perfect score: 2789  
Sequence: 1 GTGATTGTTACTTGTGCTGTCTGTTACTTAACATTCAA 1553

Scoring table: BLOSUM62

|             |             |
|-------------|-------------|
| Xgapop 10.0 | Xgapext 0.5 |
| Ygapop 10.0 | Ygapext 0.5 |
| Fgapop 6.0  | Fgapext 7.0 |
| Dgapop 6.0  | Dgapext 7.0 |

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_plus.n2p.model  
-Q=/cgn2\_1/USPTO.spool/US0974506/runat\_22082003\_132709\_10731/apz.query.fasta\_1.1735  
-DB=SwissProt\_41 -QPM=fastan -SUFFIX=n2p -ISP=MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=100 -THR\_SCORE=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORMEXT=HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0974506.ecgn\_1.1.40 -runat\_22082003\_132709\_10731 -NCP=6 -ICP=3  
-NO\_MAP -LANG=ENGLISH -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1799  | 64.5        | 350    | 1     | NF3L_HUMAN  |
| 2          | 1575  | 56.5        | 350    | 1     | NF3L_MOUSE  |
| 3          | 578   | 20.7        | 292    | 1     | A36A_DROME  |
| 4          | 453.5 | 16.3        | 278    | 1     | YJUC_SCHPO  |
| 5          | 384.5 | 13.8        | 288    | 1     | NIF3_YEAST  |
| 6          | 297   | 10.6        | 262    | 1     | YK04_CLOPE  |
| 7          | 285.5 | 10.2        | 373    | 1     | YOF0_BACSU  |
| 8          | 278   | 10.0        | 366    | 1     | YF59_STRAM  |
| 9          | 264   | 9.5         | 268    | 1     | YD03_CLOAB  |
| 10         | 263.5 | 9.4         | 372    | 1     | YD80_BACHD  |
| 11         | 256   | 9.2         | 373    | 1     | YEB8_LISIN  |
| 12         | 240   | 8.6         | 373    | 1     | YES2_LISMO  |
| 13         | 225   | 8.1         | 241    | 1     | Y705_CAMJE  |
| 14         | 216.5 | 7.8         | 379    | 1     | YM30_MYCTU  |
| 15         | 210   | 7.5         | 262    | 1     | Y931_STRPY  |
| 16         | 209.5 | 7.5         | 263    | 1     | YH16_ANASP  |
| 17         | 208   | 7.5         | 285    | 1     | YN01_STRCO  |
| 18         | 203.5 | 7.3         | 265    | 1     | YG09_STRPO  |

|    |       |     |      |   |            |                    |
|----|-------|-----|------|---|------------|--------------------|
| 19 | 199.5 | 7.2 | 385  | 1 | YM30_MYCLE | 069481 mycobacteri |
| 20 | 198   | 7.1 | 262  | 1 | Y351_UREPA | 09960 ureaplasma   |
| 21 | 193   | 6.9 | 243  | 1 | Y959_HELPU | 02832 helicobacte  |
| 22 | 191   | 6.8 | 244  | 1 | Y927_MERTU | 09637 methanococ   |
| 23 | 187.5 | 6.7 | 257  | 1 | YA73_LACIA | 025613 lactococcus |
| 24 | 181   | 6.5 | 243  | 1 | Y959_HELPU | 025613 helicobacte |
| 25 | 148   | 5.3 | 251  | 1 | Y456_MYCPV | 09890 mycoplasma   |
| 26 | 133   | 4.8 | 247  | 1 | YBGT_SALTY | 08377 salmonella   |
| 27 | 129   | 4.6 | 247  | 1 | YBGT_ECOLI | 07573 escherichia  |
| 28 | 117   | 4.2 | 248  | 1 | Y468_BORBU | 05142 borrelia bu  |
| 29 | 117   | 4.0 | 252  | 1 | YK93_VIBCH | 09347 vibrio chol  |
| 30 | 111   | 4.0 | 251  | 1 | Y183_PASMO | 09CP12 pasteurella |
| 31 | 110   | 3.9 | 248  | 1 | Y731_RALSO | 08XV60 ralsosula s |
| 32 | 107   | 3.8 | 250  | 1 | Y627_PYRHO | 058361 pyrococcus  |
| 33 | 106   | 3.8 | 252  | 1 | YBES_PSEAE | 09HYX2 pseudomonas |
| 34 | 105.5 | 3.8 | 3829 | 1 | SACS_HUMAN | 09ZJ4 homo sapien  |
| 35 | 103.5 | 3.7 | 3830 | 1 | SACS_MOUSE | 09J1C8 mus musculu |
| 36 | 103   | 3.7 | 674  | 1 | XP22_HUMAN | 043895 homo sapien |
| 37 | 102.5 | 3.7 | 1153 | 1 | A3D1_HUMAN | 014617 homo sapien |
| 38 | 102   | 3.7 | 249  | 1 | Y382_NEIMA | 09JW6 neisseria m  |
| 39 | 101   | 3.6 | 828  | 1 | SC10_ARATH | 081W5 arabidopsis  |
| 40 | 100   | 3.6 | 249  | 1 | YK54_NEITB | 09JX9 neisseria m  |
| 41 | 100   | 3.6 | 841  | 1 | HEIX_MERYA | 057742 methanococ  |
| 42 | 98    | 3.5 | 1161 | 1 | POL_SFVI   | 023074 simian foam |
| 43 | 97.5  | 3.5 | 245  | 1 | YH77_ARCFU | 028497 archaeoglob |
| 44 | 97.5  | 3.5 | 502  | 1 | GSPE_PSEAE | 000512 pseudomonas |
| 45 | 97.5  | 3.5 | 3259 | 1 | GTAN_HUMAN | 014789 homo sapien |

## ALIGNMENTS

RESULT 1  
NF3L\_HUMAN  
AC NF3L\_HUMAN STANDARD; PRT; 350 AA.  
ID Q9GZT8; Q9H2D2; Q9HC18;  
DT 28-FEB-2003 (Ref. 41, Created)  
DT 28-FEB-2003 (Ref. 41, Last sequence update)  
DT 15-SEP-2003 (Ref. 42, Last annotation update)  
DE NF3L-like protein 1 (Amyotrophic lateral sclerosis 2 chromosomal  
DE region candidate gene protein 1) (Miy018 protein) (MDS015).  
GN NF3L OR ALS2CR1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eulazaria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND SUBCELLULAR LOCATION.  
RX MEDLINE=20573864; PubMed=11124544;  
RA Tascou S., Uedelhoven J., Dikens C., Nayeria K., Engel W.,  
RA Burfelnd P.;  
RT "Isolation and characterization of a novel human gene, NIF3L, and its  
RT mouse ortholog, Nif3l1, highly conserved from bacteria to mammals.";  
RL Cytogenet. Cell Genet. 90:330-336(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21100893; PubMed=11161814;  
RA Hadano S., Yanagisawa Y., Skaug J., Fichter K., Nasir J.,  
RA Martindale D., Koop B.F., Scherer S.W., Nicholson D.W., Rouleau G.A.,  
RA Ikeda J.-E., Hayden W.R.;  
RT "Cloning and characterization of three novel genes, ALS2CR1, ALS2CR2,  
RT and ALS2CR3, in the juvenile amyotrophic lateral sclerosis (ALS2)  
RT critical region at chromosome 2q33-q34: candidate genes for ALS2.";  
RL Genomics 71:200-213(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Mao Y.M., Xie Y., Huang X.Y., Ying K., Dai J.L.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Huang C., Qian B., Yu Y., Gu W., Wang Y., Han Z., Chen Z.;  
RT "Novel genes expressed in hematopoietic stem/progenitor cells from  
RT myelodysplastic syndromes patient.";

RL submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [51]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovarian carcinoma;  
 RA Issigal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,  
 RA Nagatsuma M., Hosokura T., Kaku Y., Kodaira K., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Matsubae S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Minomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RN submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RL [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22386257; PubMed=12477937;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Rohn S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mulhany S.J.,  
 RA Bosak S.A., McEwan P.J., McKenna K.J., Malek J.A., Gunatane P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Rutherford J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerich A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE UPP0135 (NIF3) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF283538; AAC44846.1; ALT\_INIT.  
 DR EMBL: AF038949; BAB32499.1; -  
 DR EMBL: AF060513; AAG43131.1; -  
 DR EMBL: AF182416; AAG44952.1; -  
 DR EMBL: AK023378; BAB14551.1; -  
 DR EMBL: BC007654; AAH07654.1; ALT\_INIT.  
 DR Genew: HGNC:13390; NIF3L.  
 DR InterPro: IPR002678; DUF34.  
 DR Pfam: PF01784; NIF3; 1.  
 DR TIGRfams: TIGR00486; TIGR00486; 1.  
 DR CONFLICT 77 82 TWNTWK -> NLEHMKR (IN REF. 4).  
 FT  
 SQ SEQUENCE 350 AA: 38983 MW: 81FLASAD35B35ED7 CRC64;  
 Alignment Scores:  
 Pred. NO.: 1.37e-143 Length: 350  
 Score: 1799.000 Matches: 350  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 64.50% Indels: 0  
 DB: 1 Gaps: 0

| Db | 1    | Metasp    | leu       | lys     | Ala    | leu    | Ser    | Ser      | Leu     | Asn  | Pro | Phe | Ala  | Ser | Phe | Ala | Gl  | 20  |     |     |      |      |
|----|------|-----------|-----------|---------|--------|--------|--------|----------|---------|------|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|------|------|
| QY | 305  | AGTTGGACA | CAATGTTGG | ATTACTG | TGGTGA | ACCAAC | AGCCAC | CACTACTG | TAAATAC | ACTC | 364 |     |      |     |     |     |     |     |     |     |      |      |
| Db | 21   | Ser       | Trp       | Asp     | Asn    | Val    | Gly    | Leu      | Leu     | Val  | Glu | Pro | Ser  | Pro | His | Thr | Val | Asn | Thr | Leu | 40   |      |
| QY | 365  | TTTCTGAC  | CAATG     | ACTG    | TAC    | GAG    | AGT    | GTG      | TGA     | AGG  | AGG | CGC | CAAA | AG  | AG  | AG  | AG  | AG  | AG  | AG  | 424  |      |
| Db | 41   | Phe       | Leu       | Thr     | Asn    | Pro    | Leu    | Thr      | Glu     | Glu  | Val | Met | Cu   | Glu | Val | Leu | Glu | Val | Leu | Asp | 60   |      |
| QY | 425  | CTCATTTCT | CTAC      | ATCCG   | CCCTAT | CTCCG  | ACCA   | CTG      | AG      | AG   | AG  | AG  | AG   | AG  | AG  | AG  | AG  | AG  | AG  | AG  | 484  |      |
| Db | 61   | Leu       | Leu       | Ser     | Thr    | Val    | Pro    | Pro      | Thr     | Leu  | Pro | Leu | Pro  | Met | Val | Ser | Val | Leu | Thr | Asn | Thr  | 80   |
| QY | 485  | TGG       | AG        | AG      | AG     | AG     | AG     | AG       | AG      | AG   | AG  | AG  | AG   | AG  | AG  | AG  | AG  | AG  | AG  | AG  | 544  |      |
| Db | 81   | Trp       | Val       | Ser     | Glu    | Arg    | Leu    | Val      | Leu     | Arg  | Ala | Leu | Glu  | Asn | Arg | Val | Gly | Leu | Ser | Pro | His  | 100  |
| QY | 545  | AC        | AG        | CC      | CA     | TGA    | TGG    | TGG      | CG      | CC   | AG  | GG  | GG   | GG  | GG  | GG  | GG  | GG  | GG  | GG  | GG   | 604  |
| Db | 101  | Thr       | Leu       | Val     | Pro    | Ala    | Ala    | Pro      | Glu     | Glu  | Val | Asn | Asn  | Pro | Leu | Val | Ala | Val | Ser | Gly | Leu  | 120  |
| QY | 605  | TGT       | ACT       | CTC     | CA     | AG     | CC     | CA       | TAC     | ATC  | CTT | CC  | CA   | AG  | CC  | CTC | CA  | CTC | CA  | CTC | CA   | 664  |
| Db | 121  | Cys       | Thr       | Ser     | Thr    | Arg    | Pro    | Leu      | His     | Pro  | Ser | Thr | Val  | Ala | Pro | Asn | Val | Pro | Thr | Glu | Val  | 140  |
| QY | 665  | CG        | AG        | TAC     | AG     | AA     | TTC    | CA       | AG      | CC   | CA  | AG  | CC   | CTG | CA  | AA  | AG  | CTG | CTC | GC  | AG   | 724  |
| Db | 141  | Arg       | Val       | Glu     | Pro    | Phe    | Asn    | Val      | Asn     | Val  | Thr | Glu | Asn  | Pro | Val | Ser | Val | Ala | Val | Val | 160  |      |
| QY | 725  | GGA       | ATT       | G       | AG     | CG     | GT     | TTT      | CTG     | CA   | CTT | CTT | CTT  | CTG | TAC | AG  | AG  | AG  | AG  | AG  | 784  |      |
| Db | 161  | Gly       | Leu       | Ser     | Glu    | Val    | Ser    | Val      | Thr     | Ser  | Phe | Ser | Ala  | Arg | Thr | Gly | Asn | Glu | Glu | Thr | 180  |      |
| QY | 785  | CG        | GA        | T       | T      | A      | T      | C        | G       | A    | T   | T   | G    | A   | T   | T   | G   | A   | T   | T   | T    | 844  |
| Db | 181  | Arg       | Leu       | Asn     | Leu    | Asn    | Cys    | Thr      | Glu     | Val  | Ala | Leu | Met  | Glu | Val | Ala | Asp | Phe | Leu | Ser | Arg  | 200  |
| QY | 845  | AAC       | AA        | CA      | CA     | CA     | CTT    | TAT      | C       | AG   | AG  | CG  | AA   | ATT | G   | T   | C   | A   | CTG | G   | AG   | 904  |
| Db | 201  | Asn       | Val       | Ser     | Glu    | Leu    | Val    | Glu      | Val     | Thr  | Glu | Leu | Ser  | Leu | Glu | Val | Pro | Leu | Leu | His | 220  |      |
| QY | 905  | ACT       | G         | GA      | T      | T      | G      | AG       | CG      | GT   | TAT | G   | CA   | CTC | CA  | AG  | AG  | AG  | AG  | AG  | 964  |      |
| Db | 221  | Thr       | Glu       | Met     | Glu    | Arg    | Leu    | Cys      | Thr     | Leu  | Asp | Glu | Ser  | Val | Ser | Leu | Ala | Thr | Met | Leu | 240  |      |
| QY | 965  | CG        | AA        | T       | T      | A      | A      | A        | A       | A    | A   | A   | A    | A   | A   | A   | A   | A   | A   | A   | 1024 |      |
| Db | 241  | Arg       | Leu       | Ser     | Val    | Glu    | His    | Leu      | Ser     | His  | Thr | Leu | Glu  | Ala | Leu | Gly | Val | Gly | Arg | Thr | 260  |      |
| QY | 1025 | TTA       | AG        | T       | C      | T      | C      | A        | A       | G    | T   | C   | G    | T   | G   | G   | C   | C   | T   | G   | T    | 1084 |
| Db | 261  | Leu       | Glu       | Ser     | Glu    | Val    | Val    | Ser      | Val     | Ala  | Leu | Cys | Ala  | Gly | Ser | Ser | Val | Leu | Glu | Val | 280  |      |
| QY | 1085 | GG        | T         | T       | G      | T      | T      | A        | G       | G    | T   | G   | A    | C   | T   | T   | T   | A   | C   | T   | T    | 1144 |
| Db |      |           |           |         |        |        |        |          |         |      |     |     |      |     |     |     |     |     |     |     |      |      |

AC Q9E080; Q9D098; 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE NIF3-Like protein 1.  
 GN NIF3L1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20573864; PubMed=1124544;  
 RA Tascou S., Uedelhoven J., Dlxkens C., Nayeria K., Engel W.,  
 RA Burfeind P.;  
 RT "Isolation and characterization of a novel human gene, NIF3L1, and its  
 RT mouse ortholog, Nif3l1, highly conserved from bacteria to mammals.";  
 RL Cytogenet. Cell Genet. 90:330-336(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.,  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant P.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,  
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombeerts P.,  
 RA Notodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz G., Whitaker C., Wilming L.,  
 RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF284439; AAG45961.1; ALT\_INIT.  
 DR EMBL: AK011670; BAB27769.1; ALT\_INIT.  
 DR MGD: MGI:1929485; NIF3L1.  
 DR InterPro: IPR002678; DUF34.  
 DR Pfam: PF01784; NIF3; 1.  
 DR TIGRfam: TIGR00486; 1.  
 DR CONFLICT 164 164 G -> S (IN REF. 2).  
 FT CONFLICT 178 178 L -> K (IN REF. 2).  
 FT CONFLICT 191 191 L -> F (IN REF. 2).  
 FT CONFLICT 195 195 L -> Q (IN REF. 2).  
 FT CONFLICT 198 198 L -> F (IN REF. 2).  
 FT CONFLICT 208 208 T -> I (IN REF. 2).  
 SQ SEQUENCE 350 AA; 38828 MW; B0FA71503FCF7086 CRC64;

## Alignment Scores:

Pred. No.: 1,03e-124 Length: 350  
 Score: 1575.00 Matches: 302  
 Percent Similarity: 93.14% Conservative: 24  
 Best Local Similarity: 86.29% Mismatches: 24

Query Match: 56.47% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-745-506-74 (1-1553) x NIF3L\_MOUSE (1-350)  
 QY 245 ATGATTTGAAAGCTCTCTTCTTCTTGAATGATTTGAAATCCCTCTGTTGCTGAG 304  
 DB 1 MetabpLeuLylAlaLeuLeuSerLeuLeuAnpPhaLlSerLeuSerhehLaGlu 20  
 QY 305 AGTTGGACATGTTGGATTACTGTTGGAACCAAGCCCAACATCTAATATACATC 364  
 DB 21 SerTPAspAsnValGlyLeuLeuValGluProSerProhIsthrValAsnThrLeu 40  
 QY 365 TTCCTGACCAATGACCTGCTGAGAGAGATGAGAGAGAGCTCTGCAAAAGGAGAGAC 424  
 DB 41 PheLeuThrAspLeuThrGluGluValMetAspGluAlaLeuLeuInsLysAlaAsp 60  
 QY 425 CTCATCTCTCTTACCATCCGCTTCTCTCCAGCCATGAAAGCCATACCTGAAACACA 484  
 DB 61 PheIleLeuSerTyrHisProProIlePheArgProMetLysHisIleThrTrpLysThr 80  
 QY 485 TGGAAAGAGCGCTGGTATCCGGGCTGTGAGAACAGAGTGGTATCTCTCTCAT 544  
 DB 81 TrpLysGluCysLeuValIleArgAlaLeuGluAsnArgValAlaValTyrSerProHis 100  
 QY 545 AAGAGCTATGATGCTGGCGCCGAGGCGCTCAACCAACTGTTGGCTAAAGGCTTGGAGCT 604  
 DB 101 ThrAlaTyrAspAlaAlaProGlnGluValAsnSerTrpLeuAlaLysGluGlyThr 120  
 QY 605 TGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGGAAACAC 664  
 DB 121 CysThrThrArgProIleHisProSerArgAlaProAspTyrProthrgLysAlaHis 140  
 QY 665 CGATGAGATTTACGTTAACTATACACCCCAAGACCTGGACAAAGTCTGTCTGAGAA 724  
 DB 141 ArgLeuGluPheSerValAsnArgSerGlnAspLeuAspLysValMetSerThrLeuArg 160  
 QY 725 GGAATGAGCGGTTTCTCTCACTCTTTTCTGAGACATGAGTAAAGAACAAACA 784  
 DB 161 GlyValGlyGlyValSerValThrSerPheProAlaArgCysAspGlyGluGlnThr 180  
 QY 785 CGGATTAATCTGAATTTACTCAGAAAGCTTTGATGACAGTGGTATTTCTTCCCG 844  
 DB 181 ArgIleSerLeuAsnCysThrGlnLysThrLeuMetGlnValLeuAlaPheLeuSerGln 200  
 QY 845 AACCAAACTTTATATGAAAGACGAAATTTCTACTGAGAAAGCTTTGCTTCTACAT 904  
 DB 201 AspArgGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluProLeuLeuHis 220  
 QY 905 ACTGGAATGGAGCGGTTATGACACACTGATGATCTGTCCTGCAACATGATGAT 964  
 DB 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaIleMetIleGlu 240  
 QY 965 CGAATTAAGAAACACCTTAATCTCATATTCCTTACGCTTGGGGTGGGAGAAC 1024  
 DB 241 ArgIleLysThrHisLeuLysLeuSerHisLeuArgLeuAlaLeuGluGlyAlaGlyThr 260  
 QY 1025 TTAGAGTCTCAAGTCAAAATCGTGGCGCTGTGCTGCTTCTGGAGACAGCTTCTGAG 1084  
 DB 261 LeuGlnSerGlnValLysValAlaAlaLysCysAlaGlySerGlyLysSerValLeuGln 280  
 QY 1085 GGGTGTGAGCTGACCTTATACCTCACAGGTGAGATCTCCATCATATCTTTGGATGCT 1144  
 DB 281 GlyValGlnAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspValLeuAspAla 300  
 QY 1145 GCTTCCCAAGGATAAATGCTCATCTCTGTGACACAGCAACATGAAAGGCTTCTT 1204  
 DB 301 AlaSerLysGlyIleAsnValIleLeuCysGlnHisSerAspThrGlnArgLysPheLeu 320  
 QY 1205 TCTGACCTTCGAGATATGCTGATTTCTCACTTGGAGAAATAAAGTATATCTCTATCA 1264  
 DB 321 SerGlnLeuGlnGluMetLeuGlnLysHisPheGlnAsnLysIleAsnIleIleLeuSer 340

1265 GAGACGTGACAGGGGACCCCTTGTGAGGGGCA 1294  
|||||  
341 G|U|T|H|A|S|P|A|G|S|P|T|O|L|E|U|A|T|Y|A|L 350

RESULT 3  
A36A.DROME STANDARD; PRT; 292 AA.

AC C9NKK57; 09VJL4; 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Anon-35F/36A protein.  
GN ANON-35F/36A OR BG:DS02740.16 OR CG4278.  
OS Drosophila melanogaster (Fruit Fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID:7227;  
[1]  
RN RP  
RP SEQUENCE FROM N.A.  
RC STRAIN-Berkeley;  
RC MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blatz R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abdl J.F., Asghar A.I., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.F., Basu A.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernick B.P., Bhandari D., Bolshakov S.,  
RA Borkov D., Botchan M.R., Bouck B.J., Brockstein P., Brotlier P.,  
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Gleris A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.P., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclel J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E.E., Spradling A.C., Stapleton M., Stokus L., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -1- SIMILARITY: BELONGS TO THE UF0135 (NIF3) FAMILY.

```
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration.  
CC The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/  
CC or send an email to license@isb-slb.ch).  
DR EMBL; AE003415; AAF65006.1; -  
DR EMBL; AE003650; AAF653525.1; ALT_INIT.  
DR Flybase; FBgn0014093; CG4278.  
DR InterPro; IPR002678; DUF34.  
DR Pfam; PF01784; NTF3_1  
DR TIGRfams; TIGR00486; TIGR00486; 1  
SQ SEQUENCE 292 AA; 32363 MW; 775D2E7721150CC CRC64;  
  
Alignment Scores:  
Pred. No.: 1,01e+40 Length: 292  
Score: 578,00 Matches: 118  
Percent Similarity: 51,44% Conservative: 61  
Best Local Similarity: 33,91% Mismatches: 81  
Query Match: 20,72% Indels: 88  
Gaps: 4  
DB: 1  
  
US-09-745-506-74 (1-1553) x A36A_DROME (1-292)  
QY 251 TTGAAGGCCTCCTTTCTTCCTTGAAATCACTTGGACCTCTCGTTGTGAGATTGG 310  
Db LeuLaIaValVallyyscgluleuclunshphealrProthrSertrpalagluylstrp 32  
QY 311 GACAATGTGGATTACTGTGTGGAACCAAGCCACCACATFACTGTAAATACACTCTTCG 370  
Db AspasnValgLyLeuLeuileglurProhlsharglulysglnlelylsyllleueu 52  
QY 371 ACCAATGACCTGACTGAGAAGTAGTAAGAGAGGCTGTCAAAAAGAAGCACAGCTCAT 430  
Db ThrsnshpLeuthrglnurProvalVallysglnaleucluglulaleuile 72  
QY 431 CTCCTCAACATCCGCCCTATCTTCCGACCATGAAGGCATATACCTGGACATGGAAG 490  
Db IleSerrYrhIsPrrProPrllePhelysrProleuthrArgllethrGlaserhtStrpyls 92  
QY 491 GAGGGCCCTGGTGNATCCGGGCTCTGGAGAAGAGTGGATATCTCCCTCATACAGCC 550  
Db GluArgrValValAlaIaIaValasnsprllealaLeutySerProhlsthrAla 112  
QY 551 TATATGTGGGCCCCAGCGGCTCAACAACATGTGGCTGAAGAAGGCTTGAGACTGTACC 610  
Db TrpsaplylsyrseryelglyValasnsprtleuserylshalavalasnlleler 132  
QY 611 TCACAGGCCCATATACCTCTCCAAAGCTCCCAACTACCTACAGAGGAACCCAGCAGTA 670  
Db lLeatrgProleuGlurProglu----- 139  
QY 671 GAATTCACAGTTAACTACACACCCAGACCTGGACAAAGTCATGTCTGCAGTGAAGAATT 730  
Db -----LeuGlyAlarPrologly---- 145  
QY 731 GACGGTGTTCGTGCACCTCTTTTTCTGCTAGAGACTGGTAATGAGGAACAACAGGATT 790  
Db ----- 145  
QY 791 AATCTGAATGTACTACAGAAGCTTTGATGACAGGTGTAGATTCTTCCCGAACAANA 850  
Db ----- 145  
QY 851 CAACCTTATCAGAAAGCGAAATTGTGCACCTGGAAGACCTTGTCTTACATCTGGA 910  
Db -----Thrgly 147  
QY 911 ATGGGACGGTTATGACACTGGATGAATCTGCTCCCTGGCAACATGATGTATGAATA 970  
Db serglaygrlyr-----llegluhrlysmetgluleuSergerlinalValajduiserleu 165  
148 serglaygrlyr-----llegluhrlysmetgluleuSergerlinalValajduiserleu 165
```



QY 953 ACCATGATGATGCAATGAAAGACACCTAAACTATCTCATTTGCTTAGCCCTGGG 1012  
 Db 163 GluLeuValGlnAlaLysGluLeuThrGlyLeuValGlnValAlaPro 182  
 QY 1013 GTGGGAGAACCTTAGAGCTCAAGTCAAGTCTGGCCCTGTGCTGTGGAGC 1072  
 Db 183 AsnGly-----LeuAspSerHisIleSerLysValSerLeuCysAlaGlySerGlyGly 200  
 QY 1073 AGCGTTGTCAGAGGCTGTGACCTTACCTTCAAGCTGAGTATGCCATCATGAT 1132  
 Db 201 SerValValMetAsnThrAspAlaAspLeuYrPheThrGlyGluLeuSerHisHisGln 220  
 QY 1133 ACTTGGATGCTGCTTCCCAAGAAATGATCATCTGTGTAACACAGCACTGAA 1192  
 Db 221 ValLeuAlaIleMetAlaLysGlyIleSerValIleLeuGlyHisSerAsnThrGlu 240  
 QY 1193 CGAGCTTTCTTCTTGTGACCTT-----CGAATATGCTGAGTCT-----CACTTGGAGAT 1243  
 Db 241 ArgGlyLysLeuLysAspValMetCysGlnLysLeuAlaSerSerPheHisLysGlyGly 260  
 QY 1244 ---AAGATAAATATATTCCTATTCAGAGACTGACAGGACCTCTT 1285  
 Db 261 ValAspAlaAsnValIleValSerSerMetAspAlaAspProLeu 275

## RESULT 5

NIF3\_YEAST STANDARD; PRT; 288 AA.  
 AC P53081;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE NGC1-interacting factor 3.  
 GN NIF3 OR YGL221C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE-97435481; PubMed-9290212;  
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
 chromosome VII."  
 RL Yeast 13:1077-1090(1997).  
 CC -1 SIMILARITY: BELONGS TO THE UPE0135 (NIF3) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: Z72743; CA96937.1;  
 DR PIR: S64243;  
 DR SGP: S0003189; NIF3.  
 DR InterPro: IPR002678; DUF34.  
 DR Pfam: PF01784; NIF3; 1.  
 DR TIGRFAAS: TIGR00486; TIGR00486; 1.  
 DR SEQUENCE 288 AA; 31888 MW; B6AB6E48AAFA776A CRC64;

## Alignment Scores:

Pred. No.: 2.03e-24 Length: 288  
 Score: 384.50 Matches: 104  
 Percent Similarity: 44.78% Conservative: 59  
 Best Local Similarity: 28.57% Mismatches: 98  
 Query Match: 13.79% Indels: 103  
 Gaps: 11

US-09-745-506-74 (1-1553) x NIF3\_YEAST (1-288)

QY 251 TTGAAGGCTCTCTTCTCTCTTGAATGACTTTGCATCCCTCTGCTTGAG---AGT 307  
 Db 10 LeuAspLysLeuValArgSerIleThrLysPheYrProGlnLysTyrAlaAspLysSer 29  
 QY 308 TGGGACAAATGTTGGATTAATCTGTTGGACCAAGCCCAACACTACTGTAATATCA----- 361  
 Db 30 ThrAspAsnThrGlyLeuLeuIleAspCysSerThrAlaGlnValThrThrAlaAspAla 49  
 QY 362 -----CTTTCCTGACCAATGACTGAGTGAAGAGTATGATGAGAGAGTG 406  
 Db 50 AsnAlaLysThrLysValLeuLeuThrValAspPheThrLysSerValAlaGlnGluAla 69  
 QY 407 CTGCAAAAGAAAGCAAGACCTGATCTCTCTACCATCGCCCTATCTTCCAGCCATGAG 466  
 Db 70 ValAspAlaAsnCysAsnValIleValAlaLysThrIleProPheIleProSerThrAsn 89  
 QY 467 CGCATTAACCTGGAACACA---TGAAGAGAGCGCTGGAGATCGGAGCTGTGGAACAGA 523  
 Db 90 ArgLeuSerProHisThrAspProGlnHisGluThrAlaIleLysLeuIleGlnTyrGly 109  
 QY 524 GTGGGTATCTACTCTCTCATACAGCTATGATGCTGCGCCCAAGGCGTCAACATGG 583  
 Db 110 IleSerValTyrCysProHisThrAlaValAlaArgGlyValAsnAspThr 129  
 QY 584 TTGGCTAAGGCGTTGGAGCTTGACCTCCAGGCCCAACATACCTTCCAAAGTCCCAAC 643  
 Db 130 LeuValArgGlyLeuAsn----- 135  
 QY 644 TACCTACAGAGGAAACACCGAGTAGAATTCACGTTAATACCAACCAAGACCTGGAC 703  
 Db 135 ----- 135  
 QY 704 AAAGTATGTGTGACAGTGAAGAATGACGCTGTCTGTCACT---TCTTTTCTGCT 760  
 Db 136 -----AsnGlyGluAsnValAlaLysSerTyrAlaLeu 146  
 QY 761 AGGACTGTATGAGGAAACAAACCGATTAATGATTTGACTGACAAAGCTTTGATG 820  
 Db 147 GluThrValSerGlyGlu----- 152  
 QY 821 CAGGTGTAGATTTTCTTCCCGAACAACATTTATCAGAGACGGAATTTCTGTCA 880  
 Db 153 -----ThrAspAspLeu--- 156  
 QY 881 CTGGAAGACCTTTGCTTCTCATACCTGGAATGGACGCTTATGCACTGGAGATCT 940  
 Db 157 -----IleGlyTyrGlyArgPheValGluPheAsnLysAsp 168  
 QY 941 GTTCCCTGGCAACCATGATGATGAAATGAAAGACACCTAAACTATCTCATATTTGCG 1000  
 Db 169 IleSerLeuGluGlnIleValLysAsnValLysArgValLeuArgValProTyrValGln 188  
 QY 1001 TTAGCCCTTGGGGTGGGGAACCTTGAAGTCTCAAGTCAAGTCT-----GTGGCCCTG 1054  
 Db 189 ValAlaSerLeuAlaIleAlaProSerAlaThrPheAsnGlnLeuLysIleLeuValAlaVal 208  
 QY 1055 TGTGCTGTTCTGGGAGACAGCTTCTGCAAGGTGT-----GAGCGTAGACTTAACTCTC 1108  
 Db 209 CysAlaGlySerGlySerGlyValPheLysGlnLeuLysGluAspValAspLeuTyrTr 228  
 QY 1109 ACAGTGAAGATGCTCCATCATACTTACTTGCATGCTGCTCCCAAGGATAAATATGATC 1168  
 Db 229 ThrGlyGluMetSerHisHisGluValLeuLysTrpLysGluMetClyLysThrValIle 248  
 QY 1169 CTCTGTGAACACAGCAACTGAACAGAGCTTTCTTGTGACCTTCAAGATATGCTG--- 1225  
 Db 249 ValCysAsnHisSerAsnThrGlnArgLysPheLeuGlnAspValMetLysGlyLeuLeu 268  
 QY 1226 -----GATTTCACTTGGAGAAATGAATTAATATATTCATCTCAAGAGACTGACAGGAC 1279  
 Db 269 GlnAspGlyGlyHis-----GluValValIleValSerLysMetAspCysAsp 283  
 QY 1280 CCTCTTCAAGTG 1291





[2]  
 SEQUENCE FROM N.A.  
 RP STRAIN-168;  
 RX MEDLINE=98044033; PubMed-9398377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero L.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borries R., Boursier L., Brans A., Braun M., Bridgell S.C., Bron S.,  
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codan J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Enlian K.D., Erlingsson J., Fabre C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Heman A.,  
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya K., Jones L.,  
 RA Jolis B., Katamata D., Kaashara Y., Kleaer-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koeller P., Konigstein G., Krogh S., Kumano M.,  
 RA Kunita K., Lapidas A., Lardinois S., Lauer J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medvede C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Moore D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Potvorlik S., Prescott A.M.,  
 RA Priesean E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield C.,  
 RA Seliguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzneger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto S., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.;  
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 CC -1 SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.  
 CC -1 CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 270.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaborator  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>).  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch).  
 CC -----  
 CC EMBL; D84432; BAAL2492.1; ALT\_FRAME.  
 DR EMBL; 299116; CABL4447.1; .  
 DR PIR; A69954; A69954.  
 DR Subtilist; BG11661; YqfO.  
 DR InterPro; IPR002678; DUF34.  
 DR Pfam; PF01784; NIF3; 1.  
 DR TIGRFAMs; TIGR00486; TIGR00486; 1.  
 DR Hypothetical protein: Complete proteome.  
 FT COMPACT 265 266 KD > FC (IN REF. 1).  
 SO SEQUENCE 373 AA; 40920 MW; E320C1268DBBE9B CRC64;

|    |      |  |      |
|----|------|--|------|
| Db | 30   | glyeudintleglythrleaslyrpro-----llelysasnvalmet                    | 44   |
| QY | 368  | stgacsaatgactgactgaggaatgtagagagactgctgcAAAAGAAAGGACACTC           | 427  |
| Db | 45   | valthrleuaspvalleugluservalilleasplualalelgluylsgluvalaspreu       | 64   |
| QY | 428  | attgtctccsaacatccggccatgtctccgaacccatccgaacgscatracctggacaaatcg    | 487  |
| Db | 65   | llellelanishisproprolethephergetleuylshisllsesthrasplpro           | 84   |
| QY | 488  | AAGAGGCGCGTGGATCCGGCTGTGGAGAAACAGATCGGTATCTACTCTCCACACA            | 547  |
| Db | 85   | AlaeglarygllelgluylrscyeulshisaspriealvaltlyrAlaalanisthr          | 104  |
| QY | 548  | GCGTATAGTGTGGCGCGCGCGCTGCACACAGCTGGTGGTAAAGGCTTGGACCTGT            | 607  |
| Db | 105  | Asnleuaspvalalaasrglyglvalasnaspreuvalaglualaleugluleuser          | 124  |
| QY | 608  | ACCCGCGAGCGCCATACATCTTCCAAAGCGCCCAACTAC-----                       | 646  |
| Db | 125  | gluthrgluvalleu-----Alaprothrythrparproleuylsyleu                  | 140  |
| QY | 647  | -----CCTACAGAC-----GGAAC-----                                      | 661  |
| Db | 141  | AlavaltryvaltProlysglutrglugluvalvalargalalaleugllyAsnlaagly       | 160  |
| QY | 662  | -----CACCGAGTAAATTCACGTTTACATACACCCACAGCCTGGACAAAGTATGCT           | 715  |
| Db | 161  | Alaeglunisilegluylutyr-----serhisceysalapheserseuglyllegly         | 178  |
| QY | 716  | GCAGTGAAGAGATTAAGCGGTGTCTGTCTCACTCTTTTGTCTGGTAGACATGTAATAG         | 775  |
| Db | 179  | serhelysproleasprelyalalys-----proheliegluylvallelgluleu           | 196  |
| QY | 776  | GAA-----CAACACAGGATTAATCTGAATGTACTGTACAAAGCCTTGTATCGAGCTG          | 826  |
| Db | 197  | gluleuvalinsgluvalargleugluthrvalrpherolyssevalglulysalaval        | 216  |
| QY | 827  | GTAAGATTTCTTGTCCCGGACAAACACTTTATACAGAAGAGC-----GAAATCTGCA          | 880  |
| Db | 217  | lleasnlaemetilleylserhispro-----tyrglugluvalalatyrasplleuyp        | 235  |
| QY | 881  | CTGAGAAAGCCTTGTCTGTCAATACAGAGATGGAGCGGTTATGACACCGCATGAATCT         | 940  |
| Db | 236  | Valglugln--thrtrolaeglulysgluleugltagvala1g1ythrleuylshsngl        | 254  |
| QY | 941  | GTCTCCCGCGAACCATGATGATGATCAATAAAGACACSCATTAATCTCATATTCCG           | 1000 |
| Db | 255  | Meththrleuylsgluhrhealaleuhrhealalyshasrulsleuavaralasnaglvalarg   | 274  |
| QY | 1001 | TTAGCCCTTGGGCTGGGAGAACCTTGAAGTCTCAAGTCAAAAGTCTGGCCCTGTGCT          | 1060 |
| Db | 275  | Met-----Valgly--Aspalaasprekmetvalylslyvalalavalleugly             | 290  |
| QY | 1061 | GGTCTCGGAGACACCGCTTCTGCAGGGTGTAG-----GCTGACCTTACTCTACA             | 1111 |
| Db | 291  | gllyasprelyasnlytyrillelanishislaaleuysarglsglualasvaltlyrvalthr   | 310  |
| QY | 1112 | GGTAGAGATGTCATCATGATACATTTGGATGCTCGTCCCAAGGAATTAAGTCACTCTC         | 1172 |
| Db | 311  | gllyaspreuyltrphenisvalalalanishaspralamectleuuglyleuasnvalvalaspr | 330  |
| QY | 1172 | TGTAAACACACACACATGAACAGAGGCTTTCTTCTGCACCTTGCAGATATCGGATTTCT        | 1231 |
| Db | 331  | probllyns--tyrAlaeglulrlyllemetuylsgluylvaltlyrarglyleuthrser      | 349  |
| QY | 1232 | CACTTGGAGAAATAG-----ATAAATTTATCTATAGAGACTGACAGGACCTT               | 1288 |
| Db | 350  | MetcyasnasprlyslsRhegllyvalasnllleRhevalisergluthrasprthrAsnpro    | 369  |









```

DB: 1 Gaps: 11
US-09-745-506-74 (1-1553) x YEB89_LISIN (1-373)
QY 272 TTGATATGACTTTGGATCCCTCTCTGCTTGTGAGATTGGACATGTTGATTAAGTCTGCTG 331
DB 13 MetLysIleAlaProLysLysLeuAlaMetLysLysProLysLeuAlaVal 32
QY 332 GAACCAAGCCGACACATACCTGTAATATACACTCTTCCGACCAATGACATGAGGAGAA 391
DB 33 GlyAspLeuSerLysLysValArgLysValMetPhe---ThrLeuAspValLeuGlu 51
QY 392 GTGATGAGAGAGGTGCTGCAAAAGAGCAGACCTATCTCTGCTTACCATCCGCTATCC 451
DB 52 ValValAspGluAlaIleGluLysArgValAspLeuIleAlaIleHisIleProPheLeu 71
QY 452 TTGCGACCCATGAGAGCGGATACCTGAGACATGAGAGAGCGCCGTGATCCGGGCT 511
DB 72 TyrArgProThrGlnHisIleAspThrThrThrLysGlnGlyLysMetLysLysLeu 91
QY 512 CTGGAGAACAGAGTGGTATCTACTCTCCATACAGCTGATGAGTGGTGGCCGACGAGGC 571
DB 92 IleLysHisAspIleThrValPheAlaIleHisThrAsnLeuAspIleAlaGlnGly 111
QY 572 GTCAACAACAGTGGTGAAGGCTTGAAGCTTGTACCTCCAGGCCATCATCCTTCC 631
DB 112 ValAsnAspIleLeuAlaAspLeuHisLeuGlnAspThrThrMetIleGluGluThr 131
QY 632 ---AAGCTCCCACTACCCCTACAGNG----- 655
DB 132 TyrThrGluProTyrCysLysIleAlaValIleValProGluAsnGlnLeuGluSerVal 151
QY 656 ---GGAAACCACCGAGTAGAATTCACAGTAACTACACCCAGACCTG 700
DB 152 ArgLeuAlaLeuValAsnAsnGlyAlaGlyGlnIleGlyThrAsnTyrThrGlu----- 169
QY 701 GACAAAGTCATGCTGCGAGTGAAGGATTCAGCGGTGTTCTGCTCTCTTCTCTCT 760
DB 170 ---CysThrPheHisIleThrThrGlyIleGlySerPheLysProGlyThrAspAlaAsn 187
QY 761 AGGACTGGTATGAGGACAAACACGATATATG----- 796
DB 188 ProThrIleGlyLysGluThrLeuThrSerIleProGluValLysIleGluAlaIle 207
QY 797 ---AATGTACTCAGAGGCTTGTATG----- 820
DB 208 PheProGlnTyrLeuThrGluThrIleThrLysAlaValLysIleAlaHisProTyrGlu 227
QY 821 ---CAGGCGTAGATTTCTTCCCGGAAACAACACTTATTCAGAAACGGAATCTG 877
DB 228 GluProAlaIleAspValTyrThrLeuGluThrGlnThrTyrLysGlu----- 243
QY 878 TCACTGAGAGAGCCTTCTCTACATCTGGAATGGAGCGTATGACACATGATGAGAA 937
DB 244 ---GlyLeuGlnArgValGlyIleThrLeuProLys 253
QY 938 TCTGTCTCCTGGACACCATGATGATGAAATAAAGACACCTAAACCTATCTCATAT 997
DB 254 LysIleSerMetValSerPheIleAspLysLeuLysThrAlaPheAlaIleAspAsnVal 273
QY 998 CGCTTAGCCCTTGGGGTGGGAGAACCTTAGAGTCTCAAGTCACAAAGTGTGGCCCTGT 1057
DB 274 ArgPhe-----ValGly---AspLeuLysAlaAsnValGlnLysValAlaIleIle 289
QY 1058 GCTGATCTGGAGAGCGCTCTG-----CAGGCTGTGAGCTGACCTTATACCTC 1108
DB 290 GlyLysAspGlyAsnLysPheIleHisGlnAlaLysAlaThrGlyAlaAspValPheIle 309
QY 1109 ACAGGTGAGATGCCATCATGATTAATCTTGGATGCTGCTCCCAAGATTAATGTCATC 1168
DB 310 ThrGlyAspValTyrTyrHisThrAlaHisAspLeuAlaIleAsnLeuProThrIle 329
QY 1169 CTCGTGTGACACAGC-----AACACTGACAGAGCCTTCTTCTGACCTTCGAGATATG 1222

```

```

DB 330 AspaIaGlyHisAsnIleGluLysValMetLysGlyTyrLeuLysAsnLysMetGlu 349
QY 1223 CTGATTTCTCAGTTCGAGATAGATTAATTTATCTCATCAGACAGACGAGCCCT 1282
DB 350 GlnAlaLysIleLeuAspTyrGlnAlaGluPheIleValSerGluValAsnThrAspPro 369
QY 1283 CTTCAG 1288
DB 370 PheGln 371
RESULT 12
YES2_LISMO STANDARD: PRT; 373 AA.
AC P53434;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein lmo1452.
GN lmo1452.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxId=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1028 / Serovar 1/2c;
RA Klarsfeld A.D., Cossart P.;
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ESD-e / Serovar 1/2a;
RA MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Fringault L., Buchrieser C., Rusnok C., Anend A.,
RA Baguerio F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Cheuouan F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurepakt G.,
RA Madeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlutener T., Simoes N., Tierrez A.,
RA Varquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: U17284; AAA62502.1; -
DB EMBL: AL591979; CAC99530.1; -
DB PIR: AD1256; AD1256.
DB Listlist: lmo01452; -
DB InterPro: IPR002678; DUF34.
DB Pfam: PF01784; NIF3; 1.
DB TIGRfams: TIGR00486; TIGR00486; 1.
DB KW Hypothetical protein; Complete proteome.
DB SEQUENCE 373 AA; 41433 MW; C79AF8A87E2369A2 CRC64;

```

```

Alignment Scores:
Pred. No.: 3 35e-12 Length: 373
Score: 240.00 Matches: 88
Percent Similarity: 45.97% Conservative: 83
Best Local Similarity: 23.66% Mismatches: 155
Query Match: 8 61% Indels: 46
DB: 1 Gaps: 14

```

```

US-09-745-506-74 (1-1553) x YE52_LISMO (1-373)
OY 272 TTGAATGACTTTGATCCCTCGTTTGTGAGAGTGGACAAATGTGATTACTGGG 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 MetGluValIleAlaProLysLysLeuAlaMetGluLysProIleGluGluVal 32
OY 332 GAACCAAGCCCAACATATGATTAATATACATCTTCCATACCAATAGCCCTGATGAGAA 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 ---GlyAspLeuSerArgLysValArgLysIleMetPheThrLeuAspValLeuGlu 51
OY 392 GTGATGGAGAGAGTGTGCAAAAGAGACATCTATCTCTCCATACATCCGCTATC 451
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 52 ValAspGluAlaIleGluLysValAspIleIleLeuIleHisIleHisIleHisIle 71
OY 452 TTCGACCCATGAGGCAATACCTGAGACATGAGACAGAGAGCGCTGATCCGGGCT 511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 TyrArgProThrGlnHisIleAspThrThrThrLysGlnGlyLysMetIleLysLys 91
OY 512 CTGAGAACACAGATGGATATCTATCTCTCTCATACAGCTATGATGCTGCGCCAGGC 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 IleLysHisAspIleThrValPheAlaAlaHisThrAsnLeuAspIleAlaGlnGly 111
OY 572 GTCAACAATCTGTGGCTAAAGGCTTGAGAGCTTGACCTCCAGGCCCATACCTTCC 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 ValAsnAsnIleLeuAlaAspLeuHisLysGlnAsnThrThrMetIleGluGluThr 131
OY 632 AAAGTCCCACTACCCCTACAGAGGAGAAACCCAGATGAAATTCACGTAACTACACC 691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 TyrSerGluProTyrCysLysIleAlaValItyrValProGlu----- 145
OY 692 CAAGACCTGGACAAATCATGTCTGCAGTG-----AAGAAATTCAGCGGTG 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 AsnGluLeuGlnSerValArgLeuAlaLeuValAsnGlnValAlaGlnIleGlyThr 165
OY 740 TCTGTCACTCTTTCTCTGCTAGACTGT-----AATGAGAACAAACAGGATTAAT 793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 GluTyrThrGluCysThrPheHisThrThrGlyIleGlySerPheLysProGlyAlaAsn 185
OY 794 CTGAATTTACT-----CAGAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 AlaAsnProThrIleGlyLysAspAlaLeuThrSerValProGluVal-----Lys 203
OY 845 AACAAACAATTTATTCAGAG-----ACGAAATTCGTGCA----- 880
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 IleGluAlaIlePheProGlnItyrLeuThrGluThrIleThrLysAlaValLysIleAla 223
OY 881 -----CTGAGAGAGCTTGTCT-----CTACATACT----- 907
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 HisProTyrGluGlnProAlaIleAspValItyrThrLeuGlnMetGlnThrTyrLysGlu 243
OY 908 GGAATGGAGAGGTTATGACACATGATGATGATGATGATGATGATGATGATGATGATGAT 967
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 GlyLeuGlyArgValGlyMetLeuProLysLysGlnGlyMetValSerPheIleAspLys 263
OY 968 ATAAAGAACACCTAAACATATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1027
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 LeuLysThrAlaPheAlaIleAspAsnValArgPhe-----IleGly-----AspLeu 279
OY 1028 GAGTCTCAAGTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 LysThrThrValHisLysValAlaIleIleGlyLysAspGlyAsnLysPheIleHisGln 299
OY 1082 ---CAGGCTGTGAGGCTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 AlaLysSerThrGlyValAspValPheIleThrGlyAspValItyrThrHisThrGlyHis 319
OY 1139 GATGCTGCTTCCCAAGAAATATGATATCTCTCTGACACAGC-----AACACTGAA 1192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 AspLeuLeuAlaIleAsnLeuProThrIleAspAlaGlyHisAsnIleGluLysValMet 339
OY 1193 CGAGGCTTCTTCTTACCTTGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 LysGlyTyrLeuLysAsnLysMetGluGlnAlaLysIleLeuAspTyrGluAlaGlu 359

```

```

OY 1253 ATATACCTATACAGACTGACAGGACGACCTTCTGAG 1288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 PheIleValSerGluValAsnThrAspPropheGln 371
RESULT 13
ID Y705_CAMJE STANDARD: PRT: 241 AA.
AC 09PK2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein Cj0705.
GN Cj0705.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxId=197;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies K.M., Feltham D., Holt R.A.,
RA Jolley K., Kariyasekera A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC CC
DR EMBL: AL139076; CAB72979.1; -.
DR PIR: D81341; D81341.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3; 1.
DR TIGRPFAM: TIGR00486; TIGR00486; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 241 AA; 27786 MW; 1574F249905B895A CRC64;
Alignment Scores:
Pred. No.: 5,26e-11 Length: 241
Score: 225.00 Matches: 68
Percent Similarity: 37.01% Conservative: 63
Best Local Similarity: 19.21% Mismatches: 95
Query Match: 8.07% Indels: 128
Gaps: 9
US-09-745-506-74 (1-1553) x Y705_CAMJE (1-241)
OY 245 ATGATTTGAGAGCTCTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MetLysLeuSerGluIleTyrAsnPheLeuAspGlnLeuSerProPheAsnIleGlnGlu 20
OY 305 AGTTGGAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 SerThrAspAsnSerGlyIleLeuLeuGlyAspArgAspSerGlu-----IleSerThrVal 39
OY 365 TTCCTGACCAATGACCTACTGAGGAAGATGATGATGATGATGATGATGATGATGATGATGAT 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40 TyrLeuSerLeuAspIleAspGluAsnIleIleLysGlnAlaSerGluAsn-----Ser 57
OY 425 CTCATCTCTCTTACATCCGCTATCTTCTGACACCATGAGGCAATACCTGGAACACA 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 LeuIleIleThrHisHisProLeuIlePheLysGlyLeuLysAspLeuTyrAspLysThr 77

```

```

QY 485 TGAAGAGAGCGCTGGTGAATCCGGCTCTGAGAGAACAGATCGGATATCTACTCTCCAT 544
   :::: :::: :::: :::: ::::
Db 78 TyrProArGalAlaPheIleIleYsgIleMerIleYrIysAsnIleSerIleIleSerMetHis 97
   :::: :::: :::: :::: ::::
QY 545 ACACCCATGATGCTGCGCCCGCCAGGCGCTCAACAACATGCTTGCTGAAGGCTTGAGACT 604
   ||| ||||| ::::
Db 98 ThrAsnTyrAspLeuSer----- 103
   ::::
QY 605 TGTATCTCCAGGCCCATATCATCTTCCAAAGCTCCCACTACCTACAGAGGAAACAC 664
   ::::
Db 103 ----- 103
QY 665 CGAGTAGAATTCAGTTACTACACCCAGACCTGGACAAAGTCATGTCAGCTGAAA 724
   ||| ::::
Db 104 -----HisLeuAsnThrTyrPheThrGluGlu----- 112
   ::::
QY 725 GGAATTCAGCGGTGTTCTGTCATCTTTCTTCTGCTAGAGACTGGTAATGAGAACAAACA 784
   ::::
Db 113 -----IleLeuGlyPheLys----- 117
   ::::
QY 785 CGGATTAATCTGAATTTGACTCAGAGAGCTTTGATGAGTGTAGATTCTTCCCGG 844
   ::::
Db 117 ----- 117
QY 845 AACAAACACTTTCAGAGAACGGAATTTCTGTCATCGAGAGACCTTGCTTCTACAT 904
   ::::
Db 118 -----IleSerPheLysAspLysPheLeuIleTyr 127
   ::::
QY 905 ACTGGAATGGAGCGGTATGCAACATGATGATGTCGCTCCGTCGCAACATGATGAT 964
   ::::
Db 128 -----ValGluAsnSerMetSerPheGluAlaLeuLysAsp 139
   ::::
QY 965 CGAATAAAGACACCTAAACATCTCATATTCGCTTACGCTTGAGGCTGGGAGAAC 1024
   ::::
Db 140 TrpAlaLysLysLeuAsnLeuGlnIleLeuArg-----ValSerAspLys 155
   ::::
QY 1025 TTAGAGTCTCAAGTCAAAGTGTGCGCTGTGCTGTCTGCGAGACAGCTTTCGACG 1084
   ::::
Db 156 GlyLysLysAspIleLysArgIleAlaIleCysThrGlySerGlyLysPheLeuIleSer 175
   ::::
QY 1085 GGTGTTGAGCGCTTACTACCTGACAGATGATGATGATGATGATGATGATGATGATGAT 1144
   ::::
Db 176 LysValAspAlaAspCysPheLeuSerGlyAspPheLysTyrHisGlnAlaLeuGlnAla 195
   ::::
QY 1145 GCTTCCCAAGGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
   ::::
Db 196 LeuSerAsnGlnIleSerLeuIleAspLeuGlyHisPheGluSerGluArgTyrPheSer 215
   ::::
QY 1202 -----CTTTCGACTTCGAGATATGCTGCTGCTCCTGCTGCTGCTGCTGCTGCTGCT 1252
   ::::
Db 216 GlnCysLeuAlaLysAspLeuLysAsnLeu----- 225
   ::::
QY 1253 ATTATCTATCATGAGAGTACAGAGGAGCCCTCTTCAGGTGTA 1294
   ::::
Db 226 -----ProLeuGlnValIle 230
   ::::

```

```

RX MEDLINE-96295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN (2)
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleisemann R.D., Allard D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Bisher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE UP0135 (NIE3) FAMILY.
CC -I- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC IN POSITIONS 254 AND 289.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: Z70692; CAA94653.1;
DR EMBL: AE007073; -; NOT_ANNOTATED_CDS.
DR PIR: B70777; B70777.
DR TIGR: MT289;
DR Tuberculist; RV2230C;
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIE3; 1.
DR TIGRFAMs: TIGR00486; TIGR00486; 1.
KW Hypothetical protein, complete proteome.
SQ SEQUENCE 379 AA; 39397 MW; F4B5E2397426C3F3 CRC64;

Alignment Scores:
Pred. No.: 3,22e-10 Length: 379
Score: 216.50 Matches: 95
Percent Similarity: 44.01% Conservative: 63
Best Local Similarity: 26.46% Mismatches: 158
Query Match: 7.76% Indels: 43
Db: 1 Gaps: 15

US-09-745-506-74 (1-1553) x YM30_MYCTU (1-379)
QY 299 GCTGAGAGTTGGGACAAATGTTGATTCGTGGAGAACCAAGCCACACATCTGAAT 358
   ||||| ||||| ||||| ||||| |||||
Db 21 AlaGlnSerTyrPaspSerValGly--LeuValCysGlyAspProAspAspValAlaSp 39
   ::::
QY 359 ACACCTTCCTCCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
   ::::
Db 40 SerValThrValAlaValAlaAspAlaThrProAlaValAlaAspGlnValProGln----- 57
   ::::
QY 419 GCGAGACTCATCTCTCTCTACATCCGCTATCTTCGACACCATGAAAGCGATTAACCTGG 478
   ||| ||||| ::::
Db 58 AlaGlyLeuLeuValHisHisProLeuLeuArgGlyValAlaSpThrValAlaAla 77
   ::::
QY 479 AACACATGGAAGAGAGCGCTGATCCGGCTCTGGAACACAGAGTGGATCTACTCT 538
   ||||| ||||| ||||| |||||
Db 78 AsnThrProLysGlyValLeuValHisArgLeuIleArgThrGlyArgSerLeuPheThr 97
   ::::
QY 539 CCTCATACAGCTATGATGCTCGCCGCCCAAGGCGTCAACACATGCTTGCTGAAGGCTT 598
   ::::

```





305 AGTTGGCAATGTTGGATTACTGGTGGAAACCAAGCCOACCATACTGATTAATACATC 364  
 21 GlnuLyspValVallysGlyLeuGlnMet---GlySerLeuaspLysaspIleValGlyVal 39  
 365 TTTCCTGACCAATGACCTGCATGAGAAAGTGAATGAGAGAGTGTGCAAAAGAAAGGACAC 424  
 40 MetIleThrLeuaspIleargGluSerThrValAlaGluAlaIleLysAsnGluValAsp 59  
 425 CTCATTCTCTCTACACATCCGCGCTATCTTCGACCCCATGAAGCCCATTAACCTGGAAACA 484  
 60 LeuIleIleThrLysShsAlaProIlePheLysProLeuLysaspLeuValSerSerPro 79  
 485 TGGAAAGAGCCCTCGTGATCCGGGCTTCGAGAAACAGATGCGGTATCATCTCCCAT 544  
 80 GlnArgAsp---IleLeuLeuaspLeuValLysShsaspIleSerValTyrValSerHis 98  
 545 ACAGCCTATGATGCTGCGCCGCCAGGCGCTCAACAACATGCTGGTTGGCTAAAGGCTTGAGCT 604  
 99 ThrAsnIleaspIleValProGlyGlyLeuAsnAspTrpPheCys----- 113  
 605 TGTACCTCCAGGCCCATACATCGTTCCAAAGCTCCAACTCCAACTACAGAGGAAACAC 664  
 113 ----- 113  
 665 CGAGTAGAATTCAACGTTTAAGTACACCCAGACACTGGACAAAGTCATGCTGCAGTGAA 724  
 114 -----AspLeuLeuGlnIleLysGluAlaThr--- 122  
 725 GGAATGTAGCGGTGTTCTGTCACTCTTTTCTGTGAGACTGGTAATGAGAAACAACA 784  
 122 ----- 122  
 785 CGGATTAAATCTGAATTGTACTCAGAAAGGCTTTGATGACAGTGAGATTTCTTTCCCGG 844  
 123 -----TyrLeuSerGlu 126  
 845 AACCAACAACATTATCAGAAAGACGAAATTCTGTCACTGAGAAAGCTTTGCTCTTACAT 904  
 127 ThrLysGluGlyPhe----- 131  
 905 ACTGGAATGGGCGGTATGCACACTGTGATGAATGTCGTCGCGGCAACCATGATTCAT 964  
 132 ---GlyIleGlyArgIleGlyThrValLysGluGln---AlaLeuGluGluLeuLaser 149  
 965 CGAATTAAGAACACACTAAACTATATCTATCTGCTTAAGCCCTTGGGGTGGGGAACCC 1024  
 150 LysValLysArgValPheaspLeuAspPheArgLeu-----IleArgTyrAsp 166  
 1025 TTAAAGTCTCAA---GTCAAAAGTCTGGCCCTGTGTGCTGTCTGGGAGACAGCCTTGTG 1081  
 167 LysGluAsnProLeuIleSerLysIleAlaIleCysGlyGlySerGlyGlyLeuPheTyr 186  
 1082 CAGGCTGTT-----GAGCGTACGCTTAACTCAGAGGGAATGTCGCATCATCAT 1132  
 187 GlnaspIleValGlnLysGlyAlaAspValTyrIleThrGlyAspIleTyrTyrHisThr 206  
 1133 ACTTTGATGCTGCTTCCCAAGGAATAATGATGCATCTGTGAAACAGACAGCAACTGAA 1192  
 207 AlaGlnGluMetLeuThrGluGlyLysLeuPheAlaValAspProGlyHis---HisIleGlu 225  
 1193 CGAGGCTTCTTTCTGACCTTCGAGATATGCTGGATTTCTCACTTGGGAATAAG----- 1246  
 226 ValLeuPheThrGluLysLeuLysGluLysLeuGlnGlyTyrLysGluLysnGlyTyr 245  
 1247 ---ATAAATATATATCTATCAGAGCTCAGAGGAGCCT 1282  
 246 AspValSerIleIleIleSerSerLysLaserThrAsnPro 258

Search completed: August 22, 2003, 14:00:24  
Job time : 45.5 secs